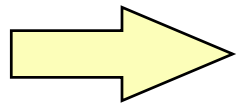# Great Lakes Environmental Data System

Jeff Sabol

AMS, Inc.

Data Management Process

April 29-30, 1999

Briefing Contents

➡️ ◆ Guiding Principles

◆ Data Management Process

◆ Lessons Learned

ams

Guiding Principles Categories

◆ Information Principles

◆ Technology Principles

◆ Relationship Principles

◆ True Multi-Media Scope
   Water, air, sediment, taxonomy, tissue, diet, meteorology

◆ Data of Documented Quality
   Three levels: Study level, Method level, Result level

◆ Extensive "Contextual" Indicators
   Ensure data longevity and ease of secondary use

Guiding Principles
Technical

◆ **User-Driven Design**
   Continuous requirements gathering and validation

◆ **Flexible and Expandable**
   Accept data from any Great Lakes monitoring project

◆ **National-Level Compatibility**
   Ease of transfer between EPA's largest water databases
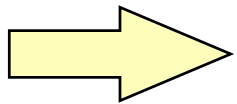
Guiding Principles
Relationship

- ◆ Vested Interest as Partners
  Multiple teams supporting GLNPO's customers

- ◆ Blend Science and Information Technology
  Knowing both sides of the "equation" at the same time

- ◆ GLNPO Owns the Database and Software
  Other groups not required for ongoing operations

Briefing Contents

✓ ◆ Guiding Principles

➡ ◆ Data Management Process

◆ Lessons Learned
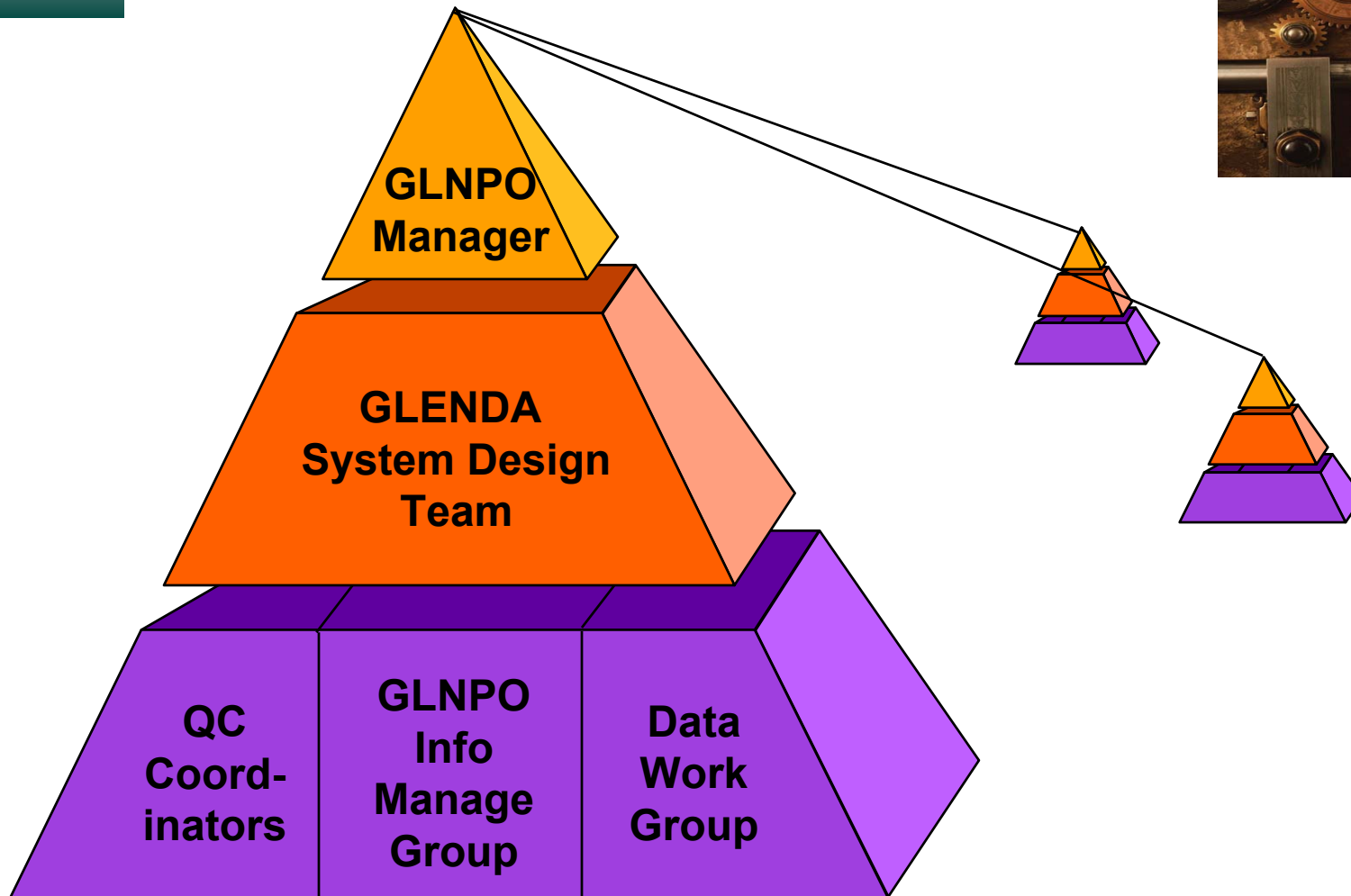
Data Management Process Components

- ◆ Governance Structure

- ◆ Foundation and Tools

- ◆ Data Administration Strategy

Data Management Process Governance Structure

- ◆ **Project Manager**
  Planning, Guidance, and Coordination Among Groups

- ◆ **GLENDA Design Team**
  Database and System Design

- ◆ **Work Groups and QC Officers**
  IT Resources, Scientific Principles, and Data Quality

ams

Data Management Process Governance Structure

GLNPO Manager

GLENDA System Design Team

QC Coord- inators

GLNPO Info Manage Group
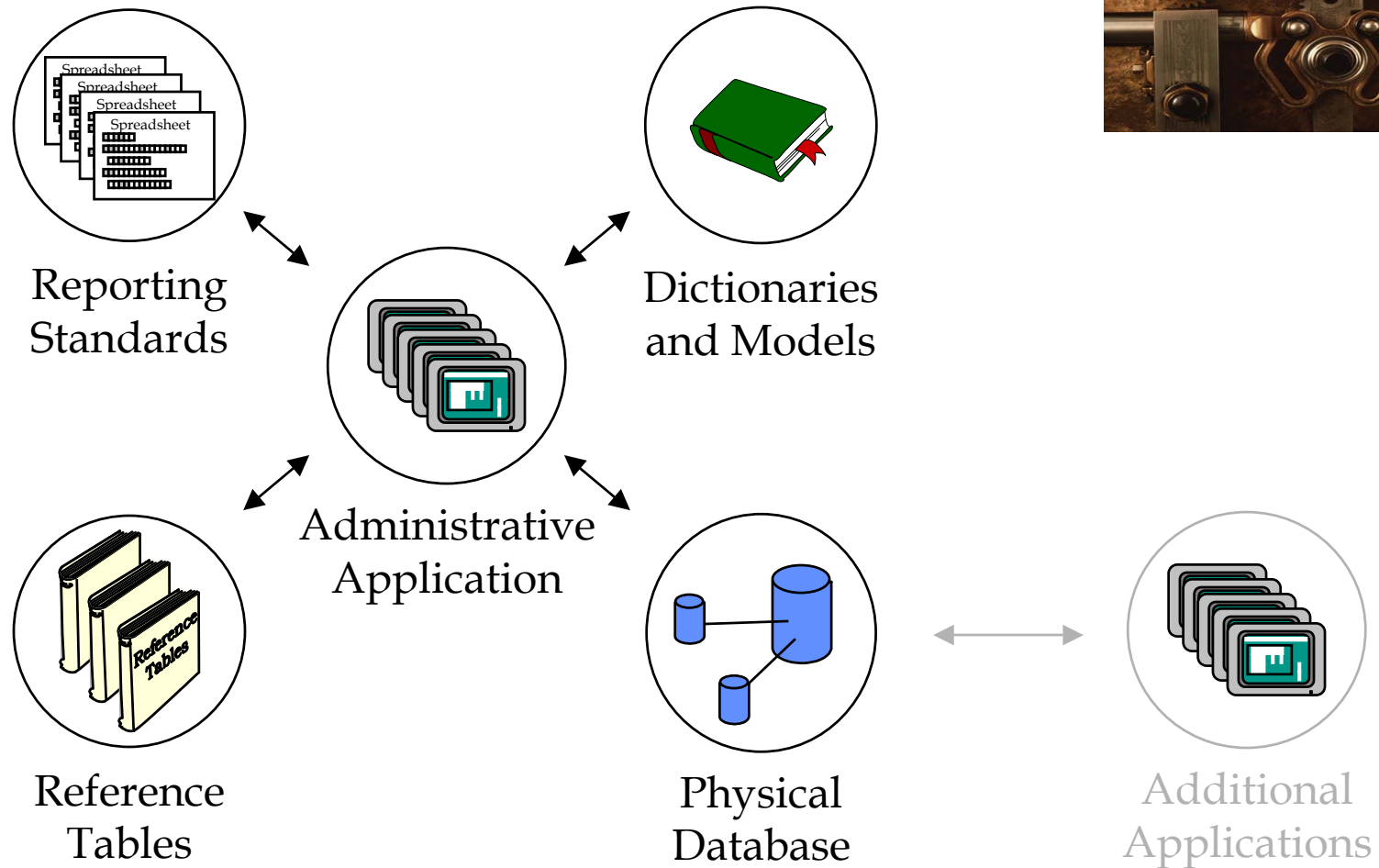
Data Work Group

ams

Data Management Process Foundation and Tools

◆ **Database, Dictionaries, and Software**
Data Uploads, Edits, Outputs …
… Security, Dictionary Viewers, Issue Manager

◆ **Reporting Standards**
Media-Specific to Reflect PI Business Areas

◆ **Reference Tables**
Valid Codes, Keyed to Reporting Standards, Distributed

Data Management Process Foundation and Tools

Reporting Standards

Dictionaries and Models

Administrative Application

Reference Tables

Physical Database

Additional Applications

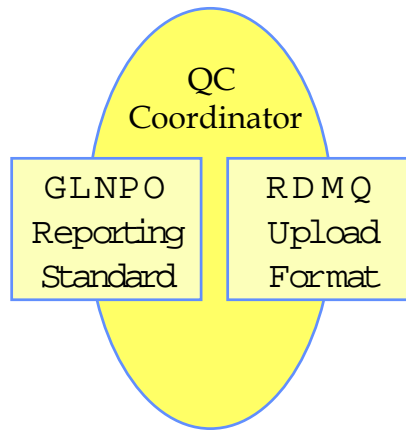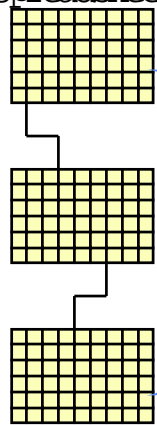ams

Data Management Process
Administrative Strategy

◆ Flow Pathways
Reference Data, Static Data, Dynamic Data

◆ Version Control and Submission Management
Internal Versions and PI Submissions

◆ Development vs. Production Environments
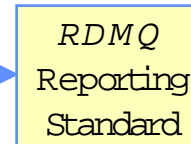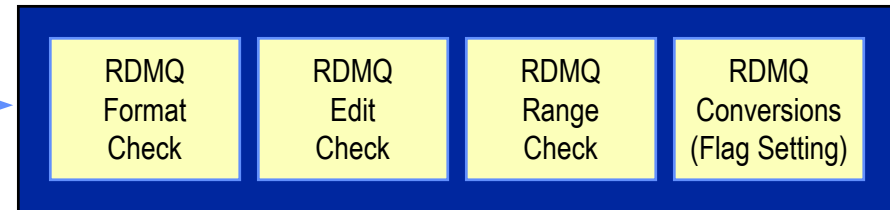*On-going Development Concurrent with Actual Use*

Data Management Process
Administrative Strategy

Individual PI
Spreadsheets

QC
Coordinator

GLNPO
Reporting
Standard

RDMQ
Upload
Format

RDMQ

| RDMQ Format Check | RDMQ Edit Check | RDMQ Range Check | RDMQ Conversions (Flag Setting) |

RDMQ
Reporting
Standard

Oracle

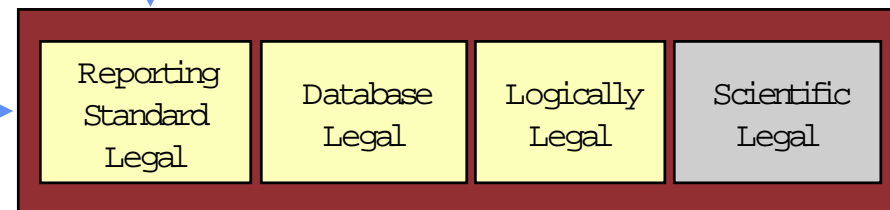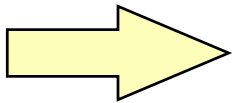| Reporting Standard Legal | Database Legal | Logically Legal | Scientific Legal |

Secondary
Users &
Public

LMMB
Modelers

ams

Briefing Contents

◆ Guiding Principles

◆ Data Management Process

◆ Lessons Learned

Lessons Learned
Organization of Topics

◆ Stumbles

◆ Success Stories

◆ Walk-Away Messages

ams

Data Management is as Sexy as Science

◆ Data Team Was Not Fully-Integrated Until Late
Planners, Project Officers, PIs, QC Coordinators worked as group for several years before data team was included

◆ Data Administration Plan Was Never Completed
Confusion on Data Pathways, Responsibilities, and Formats existed for extended period

ams

Real Expertise Pays Off

◆ Database Not Built According to Consensus Data Model
   Database built with >200 tables but only 17 were linked

◆ Original Upload Program Nonfunctional
   Consisted only of comments, No executable statements

ams

Teams are Good/Coordination is Key

◆ Original Reporting Standards Did Not Match Database
Included fields that were not in database; Omitted fields that were mandatory in database

◆ Original Reference Tables Did Not Match Database
Were not fully enumerated; Were not mutually exclusive; Contained impossible values

ams

Cross-Disciplinary Work Groups Save the Day

◆ Well-Constructed Reference Tables Now Exist
Deliberate effort by entire team resulted in corrected tables reflecting consensus and expertise of all participants

◆ Reporting Standards Developed and Tested With PIs
Deliberate effort with PIs and other pilot users resulted in usable reporting standards tailored to specific media

◆ Data Upload Process Involves Everyone
When glitches are identified, resolutions are adopted based on recommendations of all players

ams

Self-Audits Help In Course Corrections

◆ Database "Vet-Check" Identified Problems In Time
*Deficiencies were caught early enough to enable corrections
before non-recoverable impacts occurred*

◆ Semi-Annual Meetings Provided In-Progress Assistance
*Presentations and live demonstrations helped team gather
feedback, gauge reactions, and make adjustments*

◆ Peer-Review Process Lent New "Set of Eyes"
*External evaluators provided impartial assessment, distance,
and perspective (and "outside-the-box" solutions)*

ams

Ambient Monitoring "Business" Understood

◆ Database Reflects User-Specified Requirements
Supports all media, data of documented quality, EPA data standards, national database compatibility, large scale

◆ Software Reflects Administrative Needs of GLNPO
Performs data uploads/edits, data dictionary management, issue tracking, system table management, internal retrievals/transfers

◆ Data Flow Approach Reflects Real-World Experience
Working process in place, technology support can result in significant efficiency gains

ams

Lessons Learned
Walk-Away Messages

◆ **Learn from the Past But Look to the Future**
This effort built upon experience in smaller projects, mistakes are a part of the R&D process … applying the lessons later is key

◆ **Cross-Disciplinary Teams Are Critical to Success**
Real progress is made when everyone understands each other and the "business" processes they support … early coordination is best

◆ **Foundation Products are Tangible, Tested, Re-Usable Assets**
Database, Upload Software, Reporting Standards, Reference Tables, Retrieval Tools … all can be used in future projects

ams

Briefing Contents

- ◆ Guiding Principles

- ◆ Data Management Process

- ◆ Lessons Learned

ams